# 常识性知识和语篇语境对代词指认的影响*

吴 岩[1] 高约飞[1] 赵思敏[1] 王穗苹[2]

([1]东北师范大学心理学院, 长春 130024) ([2]华南师范大学心理学院，广州 510631)



**摘 要** 本研究围绕中文代词理解中，读者的常识性知识和语篇语境的作用以及作用时程这一问题展开。实验 1 首先探讨在中文阅读理解中，职业性别倾向这种常识性信息是否能对代词的指认产生早期影响。存在一致和冲突(如警卫-他；警卫-她)两个实验条件，结果在凝视时间、重读时间和总阅读时间上都发现了职业性别倾向所引发的性别冲突效应。实验 2 在实验 1 基础上增加一个代词(如：警卫-他-他；警卫-她-她)，考察更正后的信息，即文本语境内容是否可以覆盖常识的作用对代词加工产生影响。结果发现语篇语境确实可以覆盖常识性信息的作用，对代词加工产生早期影响。但是，职业性别倾向这种常识性知识仍然在代词加工的后期阶段发挥作用。考虑到代词所形成的语境较为含蓄，实验 3 中采用更明确的方式来界定职业名称的性别，例如男艺人、爸爸等，然后再出现一个代词，代词的性别始终和先前的性别描述一致，而与职业性别倾向冲突，这样也存在一致和冲突两个条件(例如：保姆-妻子-她，警卫-妻子-她)，结果发现只有更正后的语境信息对代词加工产生影响，职业性别倾向不再发挥作用。说明在中文这种高语境依赖性的语言文字中，语境可以覆盖常识性知识对代词加工产生早期作用。但是，语境作用的持续性问题会受到语篇语境中性别信息明确程度的影响。

**关键词** 代词; 语境; 常识性知识; 职业性别倾向

**分类号** B824

## 1 前言

代词是一种具有特定功能和特殊地位的词汇，代词的加工是语言理解中不可缺少的组成部分，确定代词的所指是保持句子乃至语篇连贯性的重要环节。所以对代词解决过程的探讨一直是心理语言学家所关心的问题(Garnham, 2001, a review)。作为一种代替人或事物的名称，

收稿日期：2018-6-22

*国家自然科学基金项目(31500878, 31571136)、国家哲学社会科学重点研究项目(15AZD048)资助。

通讯作者: 王穗苹, E-mail: suiping@scnu.edu.cn.

代词所携带的语义信息有限(Esaulova, Reali, & Stockhausen, 2014)。但通常情况下，代词会显示出一定的性别(中文中的“他”和“她”)和数(单复数)的区别。因此，以往很多研究在探讨代词加工过程时，都围绕着代词的性别和数量展开，通过造就先行词性别、性别倾向或数量与代词的不一致性，对代词的加工过程以及影响因素进行研究 (Canal, Garnham, & Oakhill, 2015; Duffy & Keir, 2004; Esaulova et al., 2014; Kennison & Trofe, 2003; Qiu, Chen, & Wang, 2012; Qiu, Swaab, Chen, & Wang, 2012; Xu, Jiang, & Zhou, 2013)。例如，以往研究发现，先行词的数量信息、生物性别(如：爸爸-她)都可以影响到代词的加工(Canal et al., 2015; Xu et al., 2013)；此外，先行词的性别刻板印象(如：警卫-她)也会影响到代词的指认过程(Duffy & Keir, 2004; Esaulova et al., 2014; Qiu et al., 2012)。

但是，值得注意的是，在代词指认过程中，数量或生物性别违背和性别刻板印象违背并不是一回事，前者是不可协调、不可整合的；而后者是可以协调、可以整合的。性别刻板印象是读者在生活中获得的常识性知识，并不一定是正确的。例如，警卫虽然极有可能是男性，但是并不必然是男性，也有可能是女性。因此当读者读到“警卫匆忙地拦下正要进来的出租车，因为她需要检查证件”时，虽然会产生加工困难，但是读者可以从代词“她“中推测出警卫是一名女性，从而将之前读到的信息进行整合。

可以说，在这里，代词“她”的第一次出现事实上构成了一种语境，揭示了语言材料中其所指代人物的真正性别。那么，在一个较长的语段里，读者很有可能第二次遇见代词“她”。比如，读完前面一句“警卫匆忙地拦下正要进来的出租车，因为她需要检查证件”之后，假设读者又读到以下部分“这是台里的规定，这也是她的工作，所有来访人员都必须遵守”。在这一部分里，读者会遇到第二个代词“她”。此时，将有两种对立的因素会影响到代词的加工，由于代词指代的仍是警卫，关于警卫是男性这一性别刻板印象仍有可能使读者对代词的加工发生困难；与此同时，第一个代词所提供的“对错误的职业名称性别倾向进行更正的信息”，又可以促进代词的加工。那么在这两种信息同时存在时，读者将会如何对代词进行加工呢？这事实上关系到不同类型的语义知识，如语言材料所提供的语义信息以及读者相关的常识性知识，如何影响到当前信息的加工这一问题。

以往研究关于更新后的信息是否可以取代读者头脑中固有的信息，对当前词汇的加工产生更快且持续的作用，不同的理论存在着不同的观点。来自两阶段加工模型认为，读者对当前代词的加工是分两阶段完成的，更新后的信息确实可以对当前的词汇加工产生作用，但是要在代词加工的晚期阶段才会发挥作用。例如，Kintsch的建构整合模型(Construction-Integration Model, 1988)认为，对当前信息的理解加工可以分成建构和整合两个阶段。在早期建构阶段，读者原有的一些观念首先被激活参与到当前目标加工中，而语篇的语境信息要到整合阶段才会发挥作用。Garrod和Terras的捆绑消解模型(Bonding-Resolution Model, 2000)认为阅读理解的早期阶段是一个自动化、低水平的加工过程。在此加工阶段，通过低水平的匹配机制，与当前信息有语义联系的信息首先被激活；晚期阶段为高级加工过程，在这一阶段，读者会根据语篇的语境信息对在早期阶段形成的匹配做出评价。与两阶段理论相反，另一种观点则认为阅读理解中词汇的加工是通过一个阶段就可以完成的，语境可以和常识、语义这种低水平的信息一样，对当前词汇加工产生即时性影响。例如，Hess, Foss和Carroll(1995)词汇再解释模型(Lexical Reinterpretation Model)认为读者会根据句子或语篇语境对词汇的语义进行再解释，为了确保句子或语篇信息的连贯性，只有解释后的信息才会对当前的词汇加工产生影响。

承接以上理论争端，已有一些研究对这一问题进行了探讨(Duffy & Keir, 2004; Cook & Myers, 2004; Garrod & Terras, 2000; Nieuwland & Van Berkum, 2006b)。但是，这方面的研究来自代词加工的证据较少，比较有代表性的一篇研究来自Duffy和Keir(2004)。研究者采用眼动技术，在反身代词的加工中探讨了语境和常识的作用。眼动法可以实时监控读者的阅读过程，眼动中用来测量时间的指标“注视时间”可以为不同的加工阶段提供有效的证据 (Rayner &Pollatsek, 1989)。例如，首次注视时间(first fixation duration)和凝视时间(gaze duration)通常是反映词汇早期加工的重要指标(Kliegl, Grabner, Rolfs, & Engbert, 2004)，首次注视时间指在首次通过阅读中某一兴趣区内的首个注视点的注视时间；凝视时间又称为第一遍阅读时间，指从首次注视点开始到注视点首次离开当前兴趣区之间的持续时间。重读时间(second reading time)指所有回视到当前兴趣区的注视时间之和，是反映词汇后期加工过程的重要指标(Inhoff & Liu, 1998)。

Duffy和Keir(2004)以反映词汇早期和晚期加工阶段的眼动指标凝视时间和重读时间为标准，测量了常识和语境的效应。他们给被试呈现如下材料，同时记录被试的眼动轨迹：

The electrician taught *herself/himself* a lot.

结果发现在单独给被试呈现上面材料时，冲突条件下读者对 *herself* 凝视时间和重读时间都更长。但当前面先出现“electrician is a cautious woman.”这样的信息后，读者对 *herself* 的加工无论是在凝视时间还是在重读时间上，一致和冲突条件下两者差异都不再显著。说明语言材料所提供的信息已经完全覆盖了常识性知识，对代词加工产生了即时且持续的影响，这种结果为 Hess 等(1995)提出词汇再解释模型提供了证据。

然而，除此之外，目前并没有发现其他将常识和语境结合起来，探讨两者对代词作用的文章。有单独探讨常识对代词作用的文章，如先前提及的关于先行词性别刻板印象对代词作用的研究(Duffy & Keir, 2004; Esaulova et al., 2014; Kennison & Trofe, 2003; Qiu et al., 2012); 也有单独探讨语境作用的文章(Chen, Cheung, Tang, & Wong, 2000; Nieuwland & Van Berkum, 2006a)，但将两者结合起来的研究较少。可见，关于语境和常识同时作用下，对词汇加工情况的探讨，来自代词加工的研究证据非常有限。

然而我们又不能简单地将非代词加工和代词加工混为一谈，这是因为代词加工和其他词汇，甚至于其他类型的替代词(如这个人)加工都存在区别(Crocker, Pickering, & Clifton, 2000; Esaulova et al., 2014)。Esaulova 等(2014)认为由于代词携带的语义信息有限，所以遇到代词，读者需要立即确定代词的所指，才能保证句子乃至篇章内容的连贯；而其他词汇则不同，他们含有丰富的语义信息，很多情况下不必要立即和前语境联系起来，就可以完成对当前内容的理解 (Esaulova et al., 2014; p798)。这就导致先行词乃至句子语篇语境所提供的信息在代词和其他类型词汇加工中所产生的作用不同。在 Esaulova 等研究中，研究者发现先行词所携带性别信息对代词的影响早于对其他类型替代词的作用。

此外，这方面的研究尚缺乏来自中文的研究证据，中文和英文相比，具有很多独特的性质。书面汉语的特点被概括为高度语境依赖性(Chen et al., 2000; Wang, Chen, Yang, & Mo, 2008)，表现为中文里单个汉字的意思比较模糊，在很多情况下，单字的字义必须依赖上下文才能确定。即使是意义相对较单纯的代词，如他，可能是单数的第三人称，但也可能是构

成其他词汇的一个语素，如“其他、他乡、排他”，这里的他并不指具体某个对象，也因而不需要读者去解决他所指的对象。相比之下，大部分西方的语言，如英语的词汇与词汇之间有固定的间隔，词汇的意义具有相当高的透明度，词的意义一般不会随着前后词义的改变而变化。中文的这一特点将会如何影响到语境对代词的加工呢？一种可能是，由于上下文的信息要常常被用来确定字词的意义，因而语境的作用也许会更快或更强；然而，另一种可能性是因为汉字的字词意义通常较为模糊，因而更多的加工资源将被用于确定字词的含义，从而使语境信息的利用更为延迟。总体来说，尽管不少研究者认为，相比于语法结构等因素，语义因素对成功地理解书面汉语起着相当重要的作用，然而，遗憾的是，与这种重要性相比，却鲜有研究集中去探讨不同类型的语义因素，例如语境和常识，分别是如何起作用的。

正是基于上述考虑，我们采用较为敏感的眼动技术，分三个实验探讨中文代词加工里语境与常识性知识对代词的作用和作用进程。实验 1 首先检验在中文阅读理解中，职业性别倾向这种常识性信息是否能对代词的指认产生早期影响。实验 1 的材料中会出现一个代词，按照代词的性别和职业性别倾向的关系，形成一致和冲突两个版本(例：警卫—他，警卫—她)。实验 2 的材料在实验 1 的基础上，增加一个代词，两个代词总是保持一致，同样有一致和冲突两个版本(例：警卫-他-他，警卫-她-她)。对于第二个代词来说，第一个代词所提供的性别，形成了一种与常识性知识相违背的语境。因此，可以用于考察当语境提供的信息对错误的常识性知识进行更正的情况下，语境和常识性知识如何对其后的代词加工产生影响的。由于在实验 2 里，主角的性别是通过代词确定的，这种性别提示相对比较含蓄。因此在实验 3 里，将进一步对主角的性别做出外显而明确的描述，例如男艺人、爸爸等，然后再出现一个代词，代词的性别始终和先前的性别描述一致，这样也存在一致和冲突两个版本(例如：保姆-妻子-她，警卫-妻子-她)。以进一步考察当语境所提供的性别信息清楚而明确时，代词加工又将如何受常识性知识与语境信息的影响。

## 2 实验 1

考察职业的性别倾向对代词加工的影响。根据以往研究结果，我们预期性别倾向可以影响代词加工，并在代词加工的早期阶段发挥作用。

### 2.1 研究方法

#### 2.1.1 被试

根据参考文献(Anderson, Kelley, & Maxwell, 2017; Taylor & Muller, 1996)设置先验效果量值为 0.8，采用软件 GPower3.1 版本(http://www.gpower.hhu.de/)计算出计划被试量为 24 人。实际收集被试 25 名，所有被试裸视或矫正视力正常，母语均为汉语，无阅读障碍。删除三名回答问题的正确率小于 75%的被试，有效被试 22 名，男女各半。

#### 2.1.2 实验设计和材料

采用单因素被试内设计，自变量为一致性，有两个水平，代词和职业名称的性别倾向一致或冲突。首先对职业名称的性别倾向性进行等级评价，找出具有性别倾向性的职业名称用于编写实验材料，以确保职业名称和代词确实构成一致或冲突的关系。要求 32 名不参加正式实验的某高校学生对 80 个职业名称的性别倾向性进行评定。共有 1~7 评定等级，其中一半被试的问卷里 1 代表极有可能是女性，7 代表极有可能是男性；另一半被试的问卷里评价等级正好相反。数据分析时标准统一，1 代表极有可能是男性，7 代表极有可能是女性。根据平均评价等级，选出 20 个具有强男性倾向的职业名称($M$ = 2.13; $SD$ = 0.40)和 20 个具有强女性倾向的职业名称($M$ = 5.58; $SD$ = 0.40)。独立样本 $t$ 检验结果表明两者差异显著 $t_{(38)} = 23.99$, $p<0.001$。

根据筛选出来具有性别倾向的职业名称编写实验材料，每个语篇由五句话组成，首句中出现职业名称，接下来是两个过渡句，然后是目标句，目标句中会出现目标代词，最后是结语。实验材料按照代词的性别线索和职业名称的性别倾向可以分为一致和冲突两个版本，例句如下：

一致条件：*魔术师*今天晚上有一个重要演出，可是现在头却痛得厉害，于是便走到后台休息。但*他*发现后台一片混乱，根本没有休息的地方。

冲突条件：*魔术师*今天晚上有一个重要演出，可是现在头却痛得厉害，于是便走到后台休息。但*她*发现后台一片混乱，根本没有休息的地方。

按照句子主题和两种实验条件进行拉丁方平衡，构成两个系列的实验材料，每一系列包含 40 个实验材料，每一被试只接受每个语篇的一种条件。在每一系列的实验材料中又加入

26 句填充语篇，语篇的结构和实验材料类似。所以每一系列共包含 66 个语篇。

2.1.3　实验程序

被试随机接受两个实验系列中的一个，个别施测。完成实验大约需要 40 分钟。眼动仪是 SR-Research 公司开发的 EyeLinkII，采样率为 500 次/秒，被试眼睛注视和运动情况通过头盔上两个微型红外摄像机输入电脑，我们只记录被试右眼数据。显示器屏幕与被试眼睛的距离大约是 75cm, 屏幕刷新率为 150Hz。所有材料以 34 号楷体呈现, 每个汉字大小为 0.95cm×0.95cm, 相邻汉字之间有 0.25cm 的距离, 这样每个汉字大约形成 0.73° 视角。

每段实验开始前都需要进行校准以保证被试眼动轨迹记录的精确性。每次校准都包括校对(calibration)、效度检验(validation)和漂移修正(drift correction)。在校对过程中，九个校准点(白色小圆点)会依次随机出现在屏幕中心或四周。当校准点出现时，要求被试注视该点，直至该点消失。校对之后是效度检验，仍然出现九个校准点，程序和校对一样。如果效度检验成功，则进行一次漂移修正，一个校准点会随机地出现在屏幕的中心或四周，要求被试注视它。

校准成功后方可进入正式实验。66 个语篇分 22 组呈现，每三篇为一组。具体的实验流程图如图 1 所示，每组语篇呈现前有提示信息“第 X 组”，接下来是实验材料，每个语篇以多行一屏的方式呈现在电脑屏幕上，首行空两格，所有材料均是黑底白字。每组后都有一个阅读理解题，要求被试作“是”或“否”判断，以鼓励被试仔细阅读实验材料。被试自己控制阅读速度, 读完一屏后按手柄任意键往下翻页。计算机在呈现提示信息、语篇或问题之前，都会再进行一次漂移修正以保证眼动记录的精确性。也就是屏幕左上角出现一个校准点，只有被试的注视点和校准点重合时校准点才会消失，下一屏内容才会呈现。实验之前有三组练习材料，以熟悉实验流程。

| 第 X 组 | 语篇一 | 语篇二 | 语篇三 | 判断题 |
|---|---|---|---|---|

图 1　实验流程图

2.2　结果与分析

由于过长或过短的注视点被认为不能反映阅读加工(Rayner, 1998), 因而将小于 60ms 或大于 600ms 的注视点删除(Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Qiu et al., 2012)。中文的单数人称代词由一个汉字构成，同时人称代词的词频也较高，因此，本实验发现代词的略读概率非常高, 首次略读概率在 70%以上，这一点和与汉语以往研究(Qiu et al., 2012)，以及英文单数人称代词的状况非常相似(Ehrlich & Rayner, 1983)。由于数据统计是建立在有注视点的汉字上，高略读率引发数据的大量缺失，从而导致结果的可靠性降低。为了获得更多的有效数据，研究者往往将代词与其左右一定范围的语言刺激合并为一个兴趣区来分析(Van Gompel & Majid, 2004; Qiu et al., 2012)。分析的区域包括代词区和代词后区。代词区由代词和其前后两字构成(其中 10 个左右代词前为标点符号)，代词后区由代词区后三个字组成。代词后区主要考察可能出现的延迟效应或溢出效应。

数据分析时用到的指标有首次注视时间、凝视时间、重读时间和总阅读时间。前三个指标前言中有介绍，总注视时间(total fixation duration)是指落在兴趣区的所有注视点时间的总和，该指标对较慢和较长时间的认知加工过程敏感(Kliegl et al., 2004)。因此基于此四项指标的分析结果，我们可以获知代词早期、晚期以及总体加工的情况。在以下处理中，$t_1$ 均指以被试为随机误差的 $t$ 检验值，而 $t_2$ 指以项目为随机误差的 $t$ 检验值，95%CI 为 95%置信度下的置信区间。各兴趣区内平均数和标准差见表 1。

表 1　代词区和代词后区各眼动指标的平均数和标准差(ms)

| | 代词区 | | 代词后区 | |
|---|---|---|---|---|
| | 冲突 | 一致 | 冲突 | 一致 |
| 首次注视时间 | 225(37) | 217(22) | 228(30) | 224(26) |
| 凝视时间 | 296(59) | 277(63) | 262(45) | 265(56) |
| 重读时间 | 123(104) | 99(70) | 136(99) | 121(110) |
| 总阅读时间 | 381(137) | 340(104) | 365(110) | 347(108) |

注：括号内为标准差，下同

对代词区数据分析结果表明，首次注视时间在两条件间差异不显著，$t_{1(21)}$=1.52, $p$=0.43,

$d$=0.33; $t_{2(39)}$=1.72, $p$=0.093, $d$=0.27。但和一致条件相比，冲突条件下的凝视时间更长，$t_{1(21)}$=2.25, $p$=0.036, $d$=0.50; $t_{2(39)}$=2.06, $p$=0.046, $d$=0.33; *95%CI*=1.4~36;重读时间也更长，$t_{1(21)}$=2.02, $p$=0.056, $d$=0.43; $t_{2(39)}$=1.95, $p$=0.058, $d$=0.31; *95%CI*=-0.7~48; 总阅读时间也更长，$t_{1(21)}$=3.08, $p$=0.006, $d$=0.66; $t_{2(39)}$=2.93, $p$=0.006, $d$=0.47; *95%CI*=13~67。但这种冲突效应并没有延续到代词后区，对代词后区的分析结果表明，无论是首次注视时间、凝视时间、重读时间抑或是总阅读时间，在两种条件下差异均不显著($ps$>0.2)。根据 Cohen(1988)标准，凝视时间和总阅读时间的效应量都在中等偏上水平，进一步肯定了职业的性别倾向对代词加工的影响。

这表明，在中文语篇阅读理解时，读者可以很快地将职业名称所携带的性别信息和代词联系起来，因而冲突的常识信息对代词的理解造成了阻碍。这和 Duffy 和 Keir(2004)对英文代词加工的研究结果一致；和 Esaulova 等(2014)德语代词加工的研究结果稍有不同，Esaulova 等发现职业性别倾向在晚期加工指标如总阅读时间上发挥作用。说明在不同语言中，先行词的性别倾向作用时程有所不同，本研究结果在以往研究基础上肯定了像中文这种高语境依赖的语言，代词的解决还是可以十分即时地进行。那么当文本中有信息对性别刻板印象这种常识性知识进行更正后，更正后的信息又是如何影响代词加工的？实验 2 将探讨此问题。

## 3 实验 2

实验 2 的材料中包含两个代词，两个代词总是保持一致，并且和先行词职业名称的性别倾向形成一致和冲突两个条件。因此，对于第二个代词来说，将存在两种信息影响其加工。首先，由于代词的先行词仍然是职业名称，所以职业名称的性别倾向仍会对代词的加工产生影响；其次，第一个代词所提供的性别信息也会对代词的加工产生作用。因此，如果文本中第一个代词所提供的语境信息可以更新读者头脑中关于职业名称所固有的性别倾向，对代词的加工产生作用，那么在冲突条件下，读者对第二个代词的加工将不会产生困难；而如果即使有语境信息的提示，职业的性别倾向仍然会对代词加工产生影响，那么和一致条件相比，冲突条件下第二个代词的阅读时间仍会较长。

### 3.1 研究方法

#### 3.1.1 被试

同上，计划样本量为 24 人，但是考虑到事后统计功效未达到一定标准，实际收集被试 32 人，男女各 16 人，平均年龄 20 岁，所有被试裸视或矫正视力正常，母语均为汉语，无阅读障碍。

#### 3.1.2 实验设计与材料

采用单因素被试内设计，自变量为一致性，有两个水平：代词与职业名称性别倾向一致或冲突。实验材料与实验 1 相似，同样包含五个分句，与实验 1 不同的是，首句之后紧接着出现第一目标句，该句中第一次出现与职业名称的性别倾向一致或冲突的代词，接下来是过渡句，之后是第二目标句，该句中将第二次出现代词，第二个代词与第一个代词保持一致。最后是结语。此外，实验 2 中第二个代词出现的位置和实验 1 中代词的位置大致相同。例句如下：

一致条件：*魔术师*被眼前的突发事件所惊呆了，这是*他*没有想到的事情。演出就这样被搞砸，对*他*真是不公平，当初就不应该找这样的助手。

冲突条件：*魔术师*被眼前的突发事件所惊呆了，这是*她*没有想到的事情。演出就这样被搞砸，对*她*真是不公平，当初就不应该找这样的助手。

按照句子主题和两种实验条件进行拉丁方平衡，形成两个系列的实验材料，每一系列包含 40 个实验材料，每一被试只接受每个主题材料的一种条件。在每一系列的实验材料中又加入 47 句填充材料，文本的结构和实验材料类似。每一系列共有语篇 87 个，每 3 个语篇为一组，共 29 组。

#### 3.1.3 实验步骤

实验步骤与实验 1 相同。

### 3.2 结果与分析

被试回答问题的正确率均在 81%以上，说明被试都在认真阅读实验材料。同理，删除小于 60ms 或大于 600ms 的注视点。同样，数据分析以兴趣区为单位，包括 T1(第一个代词和其前后两字，其中 10 个左右代词前为逗号)、T1+1(T1 区的后三个字组成)；T2(第二个代

词和其前后两字)和 T2+1 区(T2 区的后三个字组成)。

实验结果如表 2 所示，当读者第一次遇到和先行词职业名称相别倾向相违背的代词时，同样产生了冲突效应，表现为在 T1 区内冲突条件下凝视时间更长，$t_{1(31)}$=2.31, $p$=0.028, $d$=0.41; $t_{2(39)}$=2.46, $p$=0.018, $d$=0.40; *95%CI*=2.5~42；重读时间更长，$t_{1(31)}$=1.96, $p$=0.059, $d$=0.35; $t_{2(39)}$=1.75, $p$ =0.089, $d$=0.27; *95%CI*=-1.1~53；总阅读时间也更长，$t_{1(31)}$=3.64, $p$=0.001, $d$=0.64; $t_{2(39)}$=2.77, $p$=0.009, $d$=0.43; *95%CI*=22~78。由 Cohen(1988)关于 d 系数的标准可知，凝视时间和总阅读时间都产生了中等偏上水平的效应量，说明两条件差异较可靠。但是这样冲突效应并没有持续到 T1+1 区，在 T1+1 区，两种条件在首次注视时间、凝视时间、重读时间和总阅读时间上差异均不显著($ps$>0.1)。此结果和实验 1 结果一致，说明职业的性别倾向对代词的加工产生了早期且持续的影响。

**表 2　T1 区、T1+1 区、T2 区和 T2+1 区眼动指标的平均数和标准差(ms)**

| | T1 区 | | T1+1 区 | | T2 区 | | T2+1 区 | |
|---|---|---|---|---|---|---|---|---|
| | 冲突 | 一致 | 冲突 | 一致 | 冲突 | 一致 | 冲突 | 一致 |
| 首次注视时间 | 219(38) | 210(32) | 228(43) | 224(34) | 206(31) | 218(32) | 210(29) | 216(32) |
| 凝视时间 | 321(87) | 299(65) | 289(60) | 297(81) | 286(68) | 293(76) | 260(55) | 235(42) |
| 重读时间 | 169(99) | 143(99) | 172(120) | 157(91) | 134(86) | 114(84) | 128(110) | 113(84) |
| 总阅读时间 | 455(131) | 405(125) | 440(121) | 427(95) | 377(114) | 366(119) | 359(109) | 317(87) |

但当读者再次遇到代词时(T2 区)，冲突条件下的首次注视时间反而变小 $t_{1(31)}$=2.70, $p$=0.007, $d$=0.46; $t_{2(39)}$=2.82, $p$=0.011; $d$=0.47; *95%CI*=3.0~22。凝视时间在两条件间不再存在差异。但是冲突条件下的重读加工时间比一致条件下更长 $t_{1(31)}$=2.27, $p$=0.031, $d$=0.41; $t_{2(39)}$=1.88, $p$=0.068; $d$=0.30; *95%CI*=1.9~37。此外，在 T2+1 区，和一致条件相比，冲突条件下的凝视时间更长，$t_{1(31)}$=3.22, $p$=0.003, $d$=0.75; $t_{2(39)}$=2.87, $p$=0.007; $d$=0.44, 95%*CI*=9.1~40；总阅读时间也更长，$t_{1(31)}$=2.62, $p$=0.013, $d$=0.46; $t_{2(39)}$=2.74, $p$=0.009, $d$=0.43; 95%*CI*=9.2~73。此结果说明第一个代词所形成的语境信息对第二个代词的加工产生早期影响，导致冲突条件下第二个代词的首次阅读时间和凝视时间都变短，甚至短于一致条件。但是随着加工的进程，在后期加工阶段中，职业性别倾向这种常识性知识仍然对代词的加工产生了作用，表现在 T2 区冲突条件下的重读时间；T2+1 区凝视时间和总阅读时间更长。

可见，虽然实验 2 中第二个代词与实验 1 中代词的位置基本相同，但第一个代词的存在造成了两种不同的结果。实验 1 中职业性别倾向对代词加工产生了早期且持续的作用。实验 2 中职业性别倾向的作用被延后，在词汇后期的加工阶段或更漫长的认知加工过程中发挥作用。这是因为读者注意到了第一个代词所提供的性别信息，对职业名称的性别倾向进行了重新加工，从而促进了第二个代词的早期加工。但是，意外的是，反映早期加工的首次注视时间在冲突条件下反而变得更短，我们认为此结果和读者在阅读过程中产生的预期有关。因为文本中前后两个代词始终保持一致，那么当读者加工第一个代词时，其所提供的性别信息和职业名称的冲突性，使读者意识到之前关于职业名称性别倾向的认知在当前文本中可能是不恰当的，因而在头脑中保持对职业性别倾向的预期。而一致条件下，第一个代词的非冲突性并没有使读者形成一定预期，因此在遇到第二个代词时，反映最初加工的首次注视时间在冲突条件下反而更短。但是，这种预期效应很快就消失了，反映在凝视时间上两条件间不再存在差异。

然而，值得注意的是，在再次或者更漫长的认知加工中，职业性别倾向仍然可以发挥作用，可能原因在于，读者通过单一的代词并不能完全更正读者关于先行词的性别倾向的固有观念，因而在后继的加工中，职业名称的性别倾向仍然在发挥作用。可见在实验 2 中，读者对第一个代词所提供的性别信息的加工比较曲折，最初形成强烈的预期，然后认可文本中提供的性别，但在更漫长的认知加工中，这种纠结再一次出现，职业的性别倾向仍然在发挥作用。那么当文本中通过“男艺人、父亲”等更明朗的方式揭示先行词的性别时，在漫长的认知加工中，读者是否仍然会产生纠结，职业的性别倾向是否仍然可以影响到代词的加工呢？这是实验 3 所关心的问题。

## 4 实验 3

在实验 3 里，将进一步对职业名称的性别做出外显而明确的描述，例如男艺人、爸爸等，然后再出现一个代词。代词的性别始终和先前的性别描述一致，而与职业性别倾向冲突。如果文本中外显的性别描述所提供的语境信息可以完全更新读者头脑中关于职业名称所固有的性别倾向，对代词的加工产生作用，那么在冲突条件下，读者对第二个代词的加工将不会

产生任何困难，反映后期加工的重读时间和总阅读时间在两条件也不会存在差异；反之，冲突条件下的重读时间或总阅读时间将会更长。

## 4.1　研究方法

### 4.1.1　被试

某高校大学生 24 名，男女各半。所有被试裸视或矫正视力正常，母语均为汉语，无阅读障碍。

### 4.1.2　实验设计和材料

实验设计同实验 2，一致条件下代词与职业名称性别倾向一致，与性别描述也一致；冲突条件下代词与职业名称性别倾向不一致，与性别描述一致。实验材料同实验 2，语料中描述性别的词汇和职业名称的距离，和实验 2 中第一个代词和职业名称的距离相似；代词和职业名称的距离，和实验 2 中第二个代词和职业名称的距离相似。此外，两实验材料总体字数基本一致，例句如下：

一致条件：*魔术师*在艺术团中表现特别突出，获得杰出*男艺人*的称号，为此团里提出表扬，这令*他*整整几天都处于兴奋状态，逢人就提起这件事情。

冲突条件：*舞蹈家*在艺术团中表现特别突出，获得杰出*男艺人*的称号，为此团里提出表扬，这令*他*整整几天都处于兴奋状态，逢人就提起这件事情。

和代词构成一致和冲突条件的两个职业名称在字数(3.35 vs. 3.35)、字平均笔画数(7.97 vs. 8.24)以及词频 [1](2.57 vs. 2.39)或平均字频[1](5.19 vs. 5.22)上差异均不显著($ps>0.30$)。为了确定实验 2 中的代词和实验 3 中的性别词汇所提供语境明确度确实不同，另收集 28 名被试，要求其对两实验材料中出现的职业名称的性别进行评价。材料样例如下：

实验 2：魔术师被眼前的突发事件所惊呆了，这是他/她没有想到的事情。

实验 3：魔术师/舞蹈家在艺术团中表现特别突出，获得杰出男艺人的称号。

评价材料在电脑屏幕上呈现，要求被试确定文本中主角性别并按 F(男)或 J(女)键进行反应，记录被试反应时和正确率。四种条件按拉丁方平衡形成四个版本，每个被试随机接受其中一

[1]词频库来自 Center for Chinese Linguistic PKU，字频库来自现代汉语单字频率列表，Modern Chinese Character Frequency List(网址为: http://lingua.mtsu.edu/chinese-computing)。词频和字频均以 10 为底取对数。

个版本的实验材料。

针对项目结果，以一致性和性别描述方式为自变量，正确率和反应时为因变量，分别进行两因素重复测量方差分析。在正确率上发现性别描述方式和一致性的交互作用 $F(1, 39)=4.15$, $p=0.04$, $\eta^2=.096$，简单效应分析发现，在两种性别描述下，一致和违背的差异均显著($ps<0.001$)。但以代词进行性别描述时两条件下的正确率差异(0.96 vs. 0.82)大于以性别词进行描述时两条件间的差异(0.95 vs. 0.86)。原因在于读者在以性别词汇描述职业名称的性别时，冲突条件下被试反应的准确率大大提升。说明性别描述确实提升了语境信息对性别界定的强度。

#### 4.1.3 实验步骤

与实验 1 相同。

### 4.2 结果与分析

被试回答问题的正确率都在 80%以上，说明被试都有认真阅读材料。分析区域包括 T1 区(性别描述词)、T1+1 区(性别描述词后三字)、T2 区(代词和其前后两字)和 T2+1 区(代词区后三字)。结果分析时剔除小于 60ms 或大于 600ms 的注视点，结果如表 3 所示。

表 3 T1 区、T1+1 区、T2 区和 T2+1 区眼动指标的平均数和标准差(ms)

| | T1 区 | | T1+1 区 | | T2 区 | | T2+1 区 | |
|---|---|---|---|---|---|---|---|---|
| | 冲突 | 一致 | 冲突 | 一致 | 冲突 | 一致 | 冲突 | 一致 |
| 首次注视时间 | 211(28) | 198(33) | 187(29) | 186(35) | 212(30) | 206(19) | 191(29) | 194(23) |
| 凝视时间 | 288(58) | 256(53) | 262(67) | 246(59) | 253(58) | 238(69) | 216(46) | 215(37) |
| 重读时间 | 173(94) | 133(87) | 154(121) | 129(91) | 79(50) | 77(54) | 91(78) | 73(68) |
| 总阅读时间 | 430(117) | 359(105) | 384(169) | 349(183) | 297(73) | 276(84) | 289(96) | 269(89) |

对 T1 区的数据分析结果表明，冲突条件下首次注视时间更长，基于被试分析显著，$t_{1(23)}=2.29$, $p=0.023$, $d=0.47$; $t_{2(39)}=1.54$, $p=0.132$, $d=0.24$, $95\%CI=1.3\sim25$；凝视时间更长，$t_{1(23)}=2.94$, $p=0.007$, $d=0.59$; $t_{2(39)}=2.10$, $p=0.042$, $d=0.33$, $95\%CI=9.6\sim55$；重读时间更长，$t_{1(23)}=3.68$, $p=0.001$, $d=0.75$; $t_{2(39)}=2.02$, $p=0.049$, $d=0.32$, $95\%CI=17\sim62$；总阅读时间也更长，$t_{1(23)}=5.39$, $p<0.001$, $d=1.11$; $t_{2(39)}=3.45$, $p=0.001$, $d=0.55$, $95\%CI=44\sim98$。对性别描述后区进行

分析，没有发现显著的差异(*ps*>0.1)。此外，代词区和代词后区的结果也不显著(*ps*>0.1)。此结果说明，当读者遇到和职业名称的性别倾向不一致的性别描述时，读者发生了加工困难。但读者很快就接受了这种性别信息，即时地更正了读者头脑中固有的关于职业的性别倾向，因此在遇到与职业名称性别倾向不一致的代词时，阅读理解过程不再发生困难。表现在两条件下，代词区和代词后区的所有阅读时间均无差异。此外，对比实验 2 和实验 3 的结果我们发现，语境的明确度确实影响到了语境在代词加工中的作用，当语境信息比较含蓄时，语境可以产生即时但非持续的作用；但当语境内容比较明朗时，语境可以产生即时且持续的作用。

## 5 讨论

本研究围绕在中文代词理解中，语篇语境与常识性知识对代词作用和作用进程展开。实验 1 主要考察职业的性别倾向这种常识性知识是否可以对代词加工产生早期影响。在此基础上，实验 2 和实验 3 分别在文本中增加一定的信息，对职业的性别倾向进行更正，考察更正后的信息对代词加工的作用。

首先，在实验 1 和 2 中，研究结果都肯定了职业的性别倾向对代词加工的作用，反映在当读者首次遇到与先行词职业名称性别倾向不一致的代词时，代词的凝视时间、重读时间和总阅读时间都更长。此外，本研究还发现先行词和代词的距离不会干扰到职业性别倾向对代词的指认过程。在实验 1 和 2 中，职业名称和首次出现的代词的距离虽然不同，但性别倾向这种常识性知识都可以很快地被激活并参与到当前代词的加工。这一研究结果和 Qiu 等(2012)的研究结果有所出入，Qiu 等发现距离能够调节先行词的性别倾向对代词的作用，当先行词和代词距离较近时，性别倾向对代词加工产生即时作用；先行词离代词较远时，性别倾向在代词加工的晚期阶段才发挥作用。两研究结果的不同可能在于语篇中先行词的可通达性不同。在 Qiu 等研究中，先行词是具有一定性别倾向的姓名，而本研究中先行词是具有性别倾向的职业名称。以姓名作为先行词时，语料可以随意编制。但是以职业名称为先行词，为了保持语篇内容的自然顺畅，语篇内容基本都是围绕各职业者所从事的行业内容展开，正如本研究中大部分语料所示，这就导致在先行词职业名称的可通达性更高。而先行词可通达性恰恰是影响代词能否完全被解决的关键因素。例如，依据 Love 和 McKoon (2011)的观点，

有时代词并不必然得到完全解决, 而是处于一种部分被解决的状态。而代词是否完全解决受先行词的可通达性以及建立连贯文本表征这两个因素的影响。依据以上观点，本研究中由于先行词类型和语料特点，先行词通达性相对较好，在没有其他信息的干扰下，代词出现时其能够即时激活。所以在先行词和代词间增加一两句话，不足以影响到职业性别倾向对代词的作用。

但是，当语境中有信息对职业名称的性别倾向进行更正时，性别倾向这种常识性知识对代词加工的作用时程会发生变化，而且这种变化和语境中性别信息的突显程度有关。在实验2 中，当以代词为性别信息更正职业名称的性别倾向时，更正后的信息发挥早期作用，冲突条件下的凝视时间和一致条件下的差异不再显著，职业性别倾向的作用被延后，反映在冲突条件下的重读时间和总阅读时间仍然更长；在实验 3 中，当采用性别描述更正职业名称的性别倾向时，职业性别倾向不再对代词加工产生作用。这样的结果很难纳入现有的理论体系。例如，根据 Hess 的词汇再加工理论(Hess et al., 1995)，一旦读者头脑中的常识被更新后，只有更新后的语境信息才能对代词加工产生影响，常识不再对代词加工产生影响。此理论虽然可以很好地解释实验 3 的结果，语境完全覆盖了常识的作用，对代词即时且持续的作用。但却不能解释实验 2 的结果，语境只是部分覆盖了常识的作用，对代词加工产生的即时但非持续性的作用。此外，本研究的结果也不能用 Kintsch (1998) 的建构整合模型和 Garrod 与 Terras (2000) 的捆绑消解模型来解释。依据这两个模型，代词的加工要分两个阶段完成，常识在代词加工的早期阶段发挥作用，语境只能在代词晚期的加工阶段才能产生作用。但在实验 2 和实验 3 中，我们都在反映早期加工的指标凝视时间上都发现了语境的作用。因此，可以肯定的是，语境可以在代词加工的早期阶段发挥作用。

可见，在常识和语境的影响下，代词是分阶段的还是一次完成的，似乎和所采用语境的明确度有关。如果语境的表现方式比较含蓄，如本研究实验 2 中的代词，便会得出词汇加工分为两阶段的结果；如果采用的语境表现方式比较明确，如本研究实验 3 中的性别描述，便得到语境较强的作用，从而导致词汇加工一次完成。另外，Nieuwland 和 Van Burkum(2006b) 来自非代词词汇加工的证据也可以验证语境明确度是影响语境效应的一个因素。研究者所采用的设计是营造出一种童话故事的氛围，从而使原来不合理的现象合理化。通过不断地重复

而形成一种语境，使读者认可这个童话故事。结果表明，当读者首次听到和常识不相符的词汇时，出现了语义违背效应(N400)。当违背的词汇第三次出现时，违背的名词没有发现 N400 效应，但其前面的动词仍表现出违背效应。但当词汇在语篇中第五次出现时，与常识相违背的名词和其前面的动词已不再产生任何效应。这似乎说明，当读者第三次听到与常识相违背的名词时，所形成的关于童话的语境还不够明确，不足以完全覆盖常识对词汇加工的作用。直到句子语境足够明确了，到了第五次出现才完全发挥作用。此结果表明语境明确度似乎是影响语境对词汇作用的重要中介变量，关于语境对词汇作用的一些理论，例如两阶段加工理论(Kintsch 的建构整合模型和 Garrod 与 Terras 的捆绑消解模型)或一阶段加工理论(Hess 词汇再解释理论)的对立，可能正是由于研究者所采用的语境明确度不同所导致的。

此外，值得注意的是，本研究两实验结果虽然在语境的总体作用时程上不尽相同，但是语境均表现出了早期作用机制，反映在实验 2 中的第二个代词区的首次注视时间在冲突条件下反而更短，凝视时间在两条件间差异不再显著；实验 3 中代词区在冲突条件下的首次注视时间和凝视时间和一致条件下差异均不显著。这样的结果和近年来采用眼动和脑电技术的研究结果(Cook & Myers, 2004; Duffy & Keir, 2004; Nieuwland & Van Berkum, 2006b; Hald, Steenbeek-Planting, & Hagoort, 2007)一致。表明在中文阅读理解中，虽然词汇的切分困难，容易造成词汇理解的困难，但正因为这种情况，中文词汇理解的语境依赖性更强，所以导致语境对代词加工的作用更快。此结果和一些对中文阅读理解时程的研究结果一致(Wang et al., 2008)，为中文阅读理解过程中，语境的即时作用提供了证据支持。

最后，值得注意的是，本研究采用语境对常识进行更新的研究范式，得到了语境可以对代词加工产生即时的影响。在这样的设计中，语境是语料中被明确告知的正确信息，而常识是读者头脑中固有的，但被告知是错误的知识。相对于错误的常识知识，正确的语境信息更强。因此，较强的语境信息和错误的常识性知识相比，能对代词加工产生即时影响也是可以理解的。如果语境和常识都是错误的信息，并且语境强度和常识强度相同的情况下，语境是否仍然可以对代词加工产生早期影响呢？这一问题值得继续探讨。

总之，本研究结果说明在中文这种高语境依赖性的语言文字中，语境可以覆盖常识性知识的作用，对代词加工产生早期的影响，这和以往中文加工中探讨语境作用时程的研究具有

一致性(Wang et al., 2008)。此外，语境的明确度可以影响语境作用的持续性问题，以往关于语篇中词汇加工的两阶段理论和一阶段理论的对立，可能原因在于各研究中所采用的语境的明确度不同。

## 参 考 文 献


Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: a method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science, 28*(11), 1547–1562.

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n+1 and n+2 previews simultaneously. *Visual Cognition, 16*(6), 697–707.

Canal, P., Garnham, A., & Oakhill, J. (2015). Beyond gender stereotypes in language comprehension: self sex-role descriptions affect the brain's potentials associated with agreement processing. *Frontiers in Psychology, 6*(9), 1953.

Chen, H. C., Cheung, H., Tang, S. L., & Wong, Y. T. (2000). Effects of antecedent order and semantic context on Chinese pronoun resolution. *Memory & Cognition, 28*(3), 427–438.

Cohen, J. (1988). The statistical power analysis for the behavioral sciences. *Journal of the American Statistical Association, 2nd*(334).

Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: the influences of context and world knowledge. *Journal of Memory & Language, 50*(3), 268–288.

Crocker, M. W., Pickering, M., & Clifton, C. (2000). Architectures and mechanisms for language processing. *Linguistic Inquiry, 17*(4), 623–662.

Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition, 32*(4), 551–559.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22*(1), 75–87.

Esaulova, Y., Reali, C., & Stockhausen, L. V. (2014). Influences of grammatical and stereotypical gender during reading: eye movements in pronominal and noun phrase anaphor resolution. *Language Cognition &*

*Neuroscience, 29*(7), 781–803.

Garnham, A. (Ed). (2001). *Mental models and the interpretation of anaphora*. Philadelphia: Psychology Press/ Taylor & Francis.

Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: bonding and resolution. *Journal of Memory & Language, 42*(4), 526–544.

Hald, L. A., Steenbeek-Planting, E. G., & Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain Research. 1146*, 210–218.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology General, 124*(1), 62–82.

Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception & Performance, 24*(1), 20–34.

Kennison, S. M., & Trofe, J. L. (2003). Comprehending pronouns: a role for word-specific gender stereotype information. *Journal of Psycholinguistic Research, 32*(3), 355–378.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review, 95*(2), 163–182.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1–2), 262–284.

Love, J., & McKoon, G. (2011). Rules of engagement: Incomplete and complete pronoun resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(4), 874–887.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006a). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research, 1118*(1), 155–167.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006b). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098–1111.

Qiu, L., Swaab, T. Y., Chen, H. C., & Wang, S. (2012). The role of gender information in pronoun resolution: Evidence from Chinese. *Plos One, 7*(5), e36156.

Qiu, L. J., Wang, S. P., & Chen, H. C. (2012). Pronoun processing during language comprehension: the effects of

distance and gender stereotype. *Acta Psychologica Sinica, 44*(10), 1279−1288.

[邱丽景, 王穗苹, 陈烜之. (2012). 阅读理解中的代词加工：先行词的距离与性别刻板印象的作用. *心理学报, 44*(10), 1279−1288.]

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K., & Pollatsek, A. (1989). The psychology of reading. Englewood Cliffs, NJ: Prentice-Hall.

Taylor, D. J., & Muller, K. E. (1996). Bias in linear model power and sample size calculations due to estimating noncentrality. *Communications in Statistics: Theory and Methods, 25*(7), 1595–1610.

Van Gompel, R. P. G., & Majid, A. (2004). Antecedent frequency effects during the processing of pronouns. *Cognition, 90*(3), 255–264.

Wang, S., Chen, H.C., Yang, J., & Mo, L. (2008). Immediacy of integration in discourse comprehension: Evidence from Chinese readers' eye movements. *Language and Cognitive Processes, 23*(2), 241−257.

Xu, X., Jiang, X., & Zhou, X. (2013). Processing biological gender and number information during Chinese pronoun resolution: ERP evidence for functional differentiation. *Brain &Cognition, 81*(2), 223−236.



# The effects of discourse context and world knowledge on pronoun resolution

WU Yan[1]; GAO Yuefei[1]; ZHAO Simin[1]; WANG Suiping[2]

([1]*School of Psychology, Northeast Normal University, Changchun 130024, China*)

([2]*School of Psychology, South China Normal University, Guangzhou 510631, China*)

**Abstract**

Pronoun resolution can play a vital role in narrative comprehension. Understanding nature of pronoun resolution can help us to learn more about the cognitive processes underlying comprehension. Studies have shown that comprehension processes will be interrupted when a

pronoun mismatches its prior context or the gender stereotype of its antecedent. This indicates that discourse context and world knowledge about gender stereotype can play an important role in pronoun resolution. Recently, researchers tried to combine these two factors together and to examine which factor is crucial to the pronoun resolution. The most controversial issue is that whether the discourse context could override the world knowledge which was told to be wrong by the passage, and exert earlier influence on the pronoun resolution. Therefore, the present study examined the effects of context and world knowledge as well as its time course on pronoun resolution with eye tracking measures.

In the Experiment 1, participants were asked to read the discourse with a personal pronoun congruent or incongruent with the gender stereotype of its antecedent, an occupation name. The results revealed that reading times (including gaze, second reading time and total reading time) increased when the gender of the pronoun mismatched with the gender stereotype of its antecedent.

In the Experiment 2, another personal pronoun indicating the gender of the antecedent would be inserted into the discourse as the prior context to update the readers' gender stereotype of the occupation name. Therefore, readers would meet two identical personal pronouns while reading the passage. The first pronoun provided the updated gender information for the second pronoun. Again, the results of the first pronoun indicated that the gender stereotype of occupation could influence pronoun processing immediately. As for the second pronoun, the complicated results showed discourse context had an early influence on resolution of pronouns, but with the processing went on, the gender stereotype of occupation continued to influence integration. However, when the first pronoun was changed into an obvious gender description in Experiment 3, the discourse context was found not only to exert an earlier effect but the effect would be continued as the only factor to influence the pronoun resolution.

The current results clearly suggest that both gender stereotype and discourse context can affect the comprehension of Chinese pronouns. However, when the discourse context updates the

gender stereotype of the antecedents, the updating information can override the world knowledge information to exert an earlier effect on pronoun resolution. But whether the effects will be continued depend on the strength of the discourse context. These findings provide evidence for the interactive model of sentence comprehension.

**Key words** pronoun; discourse context; world knowledge; occupation gender stereotype